

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-055869

(43)Date of publication of application : 20.02.2002

(51)Int.Cl.

G06F 12/00

G06F 13/00

G06F 17/30

(21)Application number : 2001-161581

(71)Applicant : LUCENT TECHNOL INC

(22)Date of filing : 30.05.2001

(72)Inventor : ONG PING-WEN

(30)Priority

Priority number : 2000 580149

Priority date : 30.05.2000

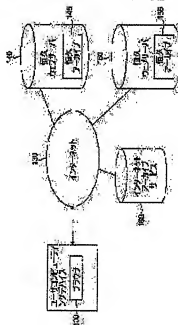
Priority country : US

## (54) INTERNET ARCHIVE SERVICE TO PROVIDE PERSISTENT ACCESS TO WEB RESOURCE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide persistent web access without necessitating direct processing of a timestamp of a URL by all web servers.

SOLUTION: A URL(Uniform Resource Locator) with the timestamp to identify web resources makes the Internet persistent by creating Internet archive service by being combined with URL redirect technology. In this invention, the persistent web access is provided without necessitating the direct processing of the timestamp in the URL by all web servers. A user inputs, for example, a URL with a date by using a web server and transmits it to the web server. The web server receives the URL (including the timestamp, a relative timestamp or a variable timestamp), retrieves a collect web page (or web pages) from the present archive or redirects the user to a predefined Internet archive site. This invention ensures that a desired material can be collated by reference with the timestamp to optional web resources.



(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-55869

(P2002-55869A)

(43) 公開日 平成14年2月20日 (2002.2.20)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード <sup>*</sup> (参考)
G 0 6 F 12/00	5 4 6	G 0 6 F 12/00	5 4 6 A 5 B 0 7 5
	5 1 7		5 1 7 5 B 0 8 2
13/00	5 2 0	13/00	5 2 0 D
17/30	2 4 0	17/30	2 4 0 B
	4 1 9		4 1 9 B

審査請求 未請求 請求項の数21 O L (全 13 頁)

(21) 出願番号 特願2001-161581 (P2001-161581)

(22) 出願日 平成13年5月30日 (2001.5.30)

(31) 優先権主張番号 0 9 / 5 8 0 1 4 9

(32) 優先日 平成12年5月30日 (2000.5.30)

(33) 優先権主張国 米国 (U S)

(71) 出願人 596092698

ルーセント テクノロジーズ インコーポ

レーテッド

アメリカ合衆国, 07974-0836 ニュージ

ヤークシー, マレイ ヒル, マウンテン ア

ヴェニュー 600

(72) 発明者 ビンウエン オン

アメリカ合衆国 07748 ニュージャーク

イ, ミドルタウン, ローレル アヴェニュー

430

(74) 代理人 100064447

弁理士 岡部 正夫 (外11名)

Fターム (参考) 5B075 N038 N020 U040

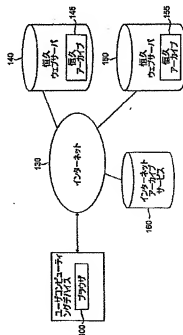
5B082 GA05 GC04

(54) 【発明の名称】 ウェブ資源に対して恒久的アクセスを提供するインターネットアーカイブサービス

## (57) 【要約】

【課題】 ウェブ資源に対して恒久的なアクセスを提供するインターネットアーカイブサービスが開示される。

【解決手段】 ウェブ資源を識別するタイムスタンプのついたURL (UniformResource Locator) は、URLリダイレクト技術と組み合わせられることで、インターネットアーカイブサービスを作成してインターネットを恒久的にしている。本発明は、すべてのウェブサーバがURLにおけるタイムスタンプを直接処理する必要なく恒久的ウェブアクセスを提供する。ユーザは、たとえば、ウェブブラウザにより日付の入ったURLを入力し、これをウェブサーバに送信する。ウェブサーバは、URL (タイムスタンプ、相対タイムスタンプまたは可変タイムスタンプを含む) を受信し、その自アーカイブから正しいウェブページ (または複数のウェブページ) を検索し、または所定のインターネットアーカイブサイトにユーザをリダイレクトする。本発明により、任意のウェブ資源に対するタイムスタンプのついた参照が所望のマテリアルを確実に照会することができる。



## 【特許請求の範囲】

【請求項1】 電子文書を提供する方法であって、前記電子文書は、それぞれタイムスタンプを有する多数のバージョンを有し、

前記電子文書の前記バージョンの1つについて、前記タイムスタンプの表示を含む要求を受信するステップと、前記電子文書の前記要求されたバージョンがアーカイブされているサーバのネットワークアドレスを送信するステップと、を含む方法。

【請求項2】 前記電子文書の前バージョンに対するポインタを、前記電子文書の前記前バージョンが略同一である場合に記憶するステップをさらに含む、請求項1記載の方法。

【請求項3】 前記電子文書を識別するアドレスは前記タイムスタンプを含む、請求項1記載の方法。

【請求項4】 前記アドレスはURL (Uniform Resource Locator) である、請求項3記載の方法。

【請求項5】 前記URL (Uniform Resource Locator) は、前記タイムスタンプを示す関連要求ヘッダを有する、請求項4記載の方法。

【請求項6】 前記要求はブラウザから受信される、請求項1記載の方法。

【請求項7】 前記送信されたネットワークアドレスは、前記ネットワークアドレスの受信機を起動して、前記電子文書の前記要求されたバージョンを、前記電子文書の前記要求されたバージョンがアーカイブされている前記サーバから入手するURLリダイレクト要求である、請求項1記載の方法。

【請求項8】 前記送信するステップは、リダイレクトURLを送信するステップをさらに含む、請求項1記載の方法。

【請求項9】 前記電子文書の前記要求されたバージョンがアーカイブされている前記サーバを識別するステップをさらに含む、請求項1記載の方法。

【請求項10】 前記識別するステップは、所定の基準を満たすサーバを識別するステップをさらに含む、請求項9記載の方法。

【請求項11】 多数のバージョンを有する電子文書を記憶するシステムであって、コンピュータ読み取り可能なコードを記憶するメモリと、

前記メモリに動作可能に接続されるプロセッサであって、

前記電子文書の前記バージョンの1つについて、前記タイムスタンプの表示を含む要求を受信し、

前記電子文書の前記要求されたバージョンがアーカイブされているサーバのネットワークアドレスを送信するように構成されるプロセッサと、を備えるシステム。

【請求項12】 前記プロセッサは、前記電子文書の前

バージョンに対するポインタを、前記電子文書の前記前バージョンが略同一である場合に記憶するようにさらに構成される、請求項11記載のシステム。

【請求項13】 前記電子文書を識別するアドレスは前記タイムスタンプを含む、請求項11記載のシステム。

【請求項14】 前記アドレスはURL (Uniform Resource Locator) である、請求項13記載のシステム。

【請求項15】 前記URL (Uniform Resource Locator) は、前記タイムスタンプを示す関連要求ヘッダを有する、請求項14記載のシステム。

【請求項16】 前記要求はブラウザにより特定される、請求項11記載のシステム。

【請求項17】 前記送信されたネットワークアドレスは、前記ネットワークアドレスの受信機を起動して、前記電子文書の前記要求されたバージョンを、前記電子文書の前記要求されたバージョンがアーカイブされている前記サーバから入手するURLリダイレクト要求である、請求項11記載のシステム。

【請求項18】 前記送信するステップは、リダイレクトURLを送信するステップをさらに含む、請求項11記載のシステム。

【請求項19】 前記プロセッサは、前記電子文書の前記要求されたバージョンがアーカイブされている前記サーバを識別するようにさらに構成される、請求項11記載のシステム。

【請求項20】 前記プロセッサは、所定の基準を満たす前記サーバを識別するようにさらに構成される、請求項19記載のシステム。

【請求項21】 電子文書にアクセスする製造物であって、

コンピュータに、前記電子文書のバージョンの1つについて、タイムスタンプの表示を含む要求を受信させ、

前記電子文書の前記要求されたバージョンがアーカイブされているサーバのネットワークアドレスを送信させるプログラムコード手段を備えるコンピュータ読み取り可能プログラムコード手段を具現させた、コンピュータ読み取り可能媒体を備える製造物。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、それぞれ本発明の譲渡人に譲渡された、1998年12月1日出願の「恒久的なウェブ資源のドメイン名を分解する方法および装置」という名称の米国特許出願第09/201,749号、1998年12月1日出願の「相対タイムスタンプを使用したウェブ資源に対する恒久的アクセス方法および装置」という名称の米国特許出願第09/201,750号、1998年12月1日出願の「ウェブ資源に対

する恒久的アクセス方法および装置」という名称の米国特許出願第09/201,751号、1998年12月1日出願の「ウェブ資源の恒久的記憶方法および装置」という名称の米国特許出願第09/201,752号、および1999年6月28日出願の「可変タイムスタンプを使用したウェブ資源に対する恒久的アクセス方法および装置」という名称の米国特許出願第09/342,408号に関連し、これらを参照して本明細書に援用する。

【0002】本発明は、インターネット資源アクセス技術に  
10 関し、特に、インターネット資源に対する恒久的アクセスを確保する方法および装置に関する。

【0003】

【従来の技術】ワールドワイドウェブ（「ウェブ」）は、膨大な量の情報を提示および頒布するための動的方法を提供する。インターネットに接続し、カリフォルニア州マウンテンビューに所在するネットスケープ・コミュニケーションズ社から市販されているネットスケープナビゲータ・コミュニケーションズ（商品名）等のブラウザを有していれば、誰でもウェブ上の情報にアクセスすることが  
20 可能である。ウェブは、ユーザに多数のメディアオプションを提供するとともに、ユーザにとってコンピュータ端末を介したその当初制限された可用性をはるかに超えて、ますます拡大している各種パーソナル電子装置においてどこでも利用可能となりつつある。しかも、このまま表示技術が進歩していくと、最終的にウェブは、従来の紙媒体に完全に取って代わる可能性がある。

【0004】

【発明が解決しようとする課題】紙媒体は、通常、関連するタイムスタンプを有し、所定時間に入手可能であった情報の決定が容易となる。たとえば、新聞記事は、新聞が発行された特定の日付を明記すれば、信頼できる参照（authoritative reference）として引用することができる。しかしながら、ウェブコンテンツの動的性質により、ウェブ文書は、一般に信頼性のある参照源ではない。現在、ウェブコンテンツは、将来のある時点で同一形式で入手可能であり、かつ同一のURL（Uniform Resource  
30 Locator）によりアドレス指定されたと確実に期待できない。ウェブサイトの中には、いくつかのアーカイブされたウェブ文書に対するアクセスを提供しているものもあるが、履歴としてのウェブ文書には、ユーザは仮にも常時かつ予測可能にアクセスすることができない。

【0005】非営利のコンピュータライブラリ提供および研究機関であるオンライン・コンピュータ・ライブラリ・センター・インク（「OCLC」）は、インターネット資源全般についてインターネットアドレスおよび  
50 エリアを管理するOCLCPURL（「Persistent Uniform Resource Loc

ator」と呼ばれるソフトウェアツールを提供している。PURLは、インターネット資源に柔軟な名付けおよび名称分解サービスを提供して、保守が最低限に抑えつつインターネット資源に対する確実かつ長期アクセスを保証する。一般に、OCLC PURLは、インターネットユーザがウェブ資源を見つける際に助けとなる。上述したように、インターネットは、常に拡大かつ変化している。URL（Uniform Resource  
Locator）が変わると、そのURLに対する過去の照会とはすべて無効になり、これによりユーザがそのインターネット資源にアクセスすることができなくなる。このような変更を管理することは面倒となること  
が多い。

【0006】URLはインターネット資源の場所を直接示す一方、PURLは、PURLを実際のURLに変換する中間分解サービスを示す。ウェブ資源がOCLCに登録され、PURLが割り当てられると、そのウェブ資源は、PURLによりアクセスされる場合がある。PURLは、資源の場所が変わっても恒久的な名称を資源に割り当てる。このように、ウェブ文書および他の資源において照会されるPURLは、ウェブ資源が移動される都度、参照を更新する必要なく恒時的に持続させておくことができる。しかしながら、OCLCにより維持されるPURL「宛先」アドレスは最新にしておかねばならない。換言すれば、OCLCは、文書が移動される都度の文書の新しいアドレスを通知されていなければならない。

【0007】非営利研究機関であるインターネット・アーカイブは、価値のある資源を保存するためにインターネットのコンテンツの定期的なスナップショットを提供している。その後、アーカイブを使用する研究のために許可された個人に対して、制限されたアクセスが提供される。アーカイブは、一般には利用可能ではなく、ウェブ資源に対する恒久的アクセスを提供しない。

【0008】

【課題を解決するための手段】全体として、恒久的アクセスをウェブ資源に対して提供するインターネットアーカイブサービスが開示される。本発明の一態様によれば、ウェブ資源を識別するタイムスタンプのついたURL（Uniform Resource Locator）は、URLリダイレクト技術と組み合わせられることで、インターネットアーカイブサービスを作り出してインターネットを恒久的なものとしている。タイムスタンプは、URLにおいて任意の適切な形式で指定されることが可能である。本発明により、ウェブは、紙媒体に非常に似た組織的で信頼性のある参照源となる。

【0009】本発明は、すべてのウェブサーバがURLのタイムスタンプを直接処理する必要なく、恒久的なウェブアクセスを提供する。一般に、ユーザは、たとえ、日付の入ったURLをウェブブラウザを用いて入

し、これをウェブサーバに送信する。開示されるウェブブラウザは、希望の日時の指定を容易にする機構をオプションで有することが可能であり、あるいは、ユーザは、ブラウザの「探索 (location)」ウィンドウにおいて表示されるURLに手でタイムスタンプを添付することができる。ウェブサーバは、URL (タイムスタンプ、相対タイムスタンプまたは可変タイムスタンプを含む) を受信し、その自アーカイブから正しいウェブページ (または複数のウェブページ) を検索し、またはインターネットアーカイブサービス160等所定のインターネットアーカイブサイトにユーザをリダイレクトする。ウェブサーバは、抽出されたURLを、選択されたタイムスタンプフォーマットにしたがって解読する。本発明により、あらゆるウェブ資源に対するタイムスタンプのついた参照が希望のマテリアルを確実に照会することができる。このように、ウェブ上で履歴の調査を行えば誰でも、いつでも有効な情報を検索することが可能となる。

#### [0010]

【発明の実施の形態】図1は、インターネットまたはワールドワイドウェブ (「ウェブ」) 環境130において1つ以上の恒久ウェブサーバ140、150からの情報にアクセスする、本発明に係るウェブブラウザ100を示す。本発明は、テキスト、音声、映像またはアニメーション文書を含むウェブ資源、すなわち電子文書に対する恒久的なアクセスを提供する。本発明の一特徴によれば、ウェブ資源を識別するURL (Uniform Resource Locator) を拡大してタイムスタンプを含むようにする。ウェブブラウザ100および恒久ウェブサーバ140、150は、付加的なタイムスタンプパラメータに対応し、これによりユーザは、厳密なターゲット日を有する任意のウェブアドレスを照会することができる。たとえば、URL "http://cnn.com?time=2+2+1997" は、1997年2月2日のCNNホームページのウェブコンテンツを指定する。このように、ウェブは、紙媒体に非常によく似た組織的で信頼性のある参照源となることができる。

【0011】図1に示すように、詳細を「インターネットアーカイブサービス」という項で後述するインターネットアーカイブサービス160は、ウェブ資源に対して恒久的なアクセスを提供する。本発明の一特徴によれば、タイムスタンプのついたURLは、URLリダイレクト技術と組み合わせられることで、インターネットアーカイブサービスを作り出してインターネットを恒久的なものとしている。一般に、ユーザは、たとえば、日付の入ったURLをブラウザ100により入力し、これを恒久ウェブサーバ140等のウェブサーバに送信する。後述するように、開示されるウェブブラウザ100は、希望の日時の指定を容易にする機構をオプションで有することが可能であり、あるいは、ユーザは、ブラウザの

「探索」ウィンドウにおいて表示されるURLに手でタイムスタンプを添付することができる。ウェブサーバ140は、URL (タイムスタンプ、相対タイムスタンプまたは可変タイムスタンプを含む) を受信し、その自アーカイブ145から正しいウェブページ (または複数のウェブページ) を検索し、またはインターネットアーカイブサービス160等所定のインターネットアーカイブサイトにユーザをリダイレクトする。インターネットアーカイブサービス160は、日付の入ったURLを用いて希望の資源 (またはその資源に対するリンク) を検索する。

【0012】(タイムスタンプフォーマットの例示) タイムスタンプは、当業者にとって明らかであるように、URL (Uniform Resource Locator) に、任意の適当な形式で含まれることが可能である。たとえば、1998年2月2日に存在したウェブページwww.lucent.comを照会するために、URLは、

http://www.lucent.com?time=2+2+1998

と表すことができる。別の変形例において、追加時間粒度 (time granularity) は、URLにおいて時刻を含むことによって表される。たとえば、1998年2月2日午後1時23分に存在したウェブページwww.lucent.comは、URLは

http://www.lucent.com?time=13+23+2+2+1998

と表すことができる。特に指定のない限り、時間帯は、ユーザのデフォルト時間帯であるものとする。上述した例示のタイムスタンプフォーマットは、共通ゲートウェイインタフェース (CGI) 検索引数である。月、日および年号 (または他の時間単位) をいずれの順序で表現してもよいことは勿論である。タイムスタンプのないURLの場合、デフォルト値は最新版となる。

【0013】さらに、相対タイムスタンプを使用してもよい。たとえば、昨日存在したウェブページwww.lucent.comを照会するために、URLは、

http://www.lucent.com?time=yesterday

と表すことができる。さらに、埋め込みハイパーリンクが相対タイムスタンプを含む場合、相対タイムスタンプは現行ウェブページが基準となる。したがって、現行ウェブページが以下の形式でURLを有し、

http://www.lucent.com?time=1998\_1\_10

当該ページが以下の形式で埋め込みハイパーリンクを含む場合、

http://www.lucent.com?time=yesterdayブラウザおよびサーバは、このURLを

http://www.lucent.com?rti  
me=1998\_1\_9  
として解読し、変換する。他の相対タイムスタンプとして、+/-指定された時間等、現行ウェブページの時間からの時間偏差を用いてもよい。たとえば、“+10D”は、現在閲覧しているウェブページの時間に10日プラスすることを示す。

【0014】(現行文書のタイムスタンプに相対して)文書の前または次にアーカイブされたバージョンを照会するためには、URLは、“next\_archive”または“previous\_archive”というラベルにより表すことができる。別の変形例において、文書の最初のまたは最新のアーカイブされたバージョンは、それぞれ“first\_archive”または“final\_archive”というラベルにより表すことができる。サーバは、アーカイブを検索して必要な文書を見つける。たとえば、現在閲覧している文書が以下の形式でURLを有し、

http://www.lucent.com/doc.html?rtime=1997\_5\_12  
7/2/97, 6/1/97, 4/1/97および3/1/97にアーカイブされた同一文書の異なるバージョンが存在する場合、これらのアーカイブの1つを照会するために、現行文書において以下のリンクが使用できる。

doc.html?rtime=next\_archive

これらの相対的なアーカイブタイムスタンプは、同一文書の異なるバージョンの間の行き来をより効率良くする。

【0015】タイムベースパラメータは、相対タイムスタンプの照会日を指定するために使用される。たとえ \*

例	意味
199(1,2)	1991または1992
199(0-9)	1990ないし1999
19(2-3)(4-5)	1924, 1925, 1934, または1935
19(20-30, 88, 90)	1920ないし1930, 1988, または1990
190.	1900ないし1909
19..	1900ないし1999
19*	19, 190ないし199, 1900ないし1999, ...

【0018】したがって、上記タイムスタンプパターン ※ するために用いられる。たとえば、は、一致するタイムスタンプを有する資源の一覧を要求※

例	意味
http://www.a.com/res.html?rtime=1998_*	1998年におけるすべてのres.htmlページ
http://www.a.com/res.html?rtime=1998_10_*	1998年10月におけるすべてのres.htmlページ
http://www.a.com/res.html?rtime=*	アーカイブされているすべてのres.htmlページ
http://www.a.com/res.html?rtime=1998_10_(2,3)	1998年10月2日および

\*ば、  
timebase=1998\_6\_11&time=+100D

は、1998年6月11日から100日後の日付を示す。同様に、

timebase=1998\_6\_11&time=next\_Monday

は、1998年6月11日以降の月曜日の日付を示す。

【0016】本発明の別の特徴によれば、タイムスタンプにおけるワイルドカードキャラクタおよび日付の範囲を使用することで、ユーザが特定のウェブ資源の日付を確信できない場合にURLにおける可変タイムスタンプを実施することが可能である。このように、サーバは、タイムスタンプパターンに一致する特定のウェブ資源の一覧を表示することが可能である。例示の実施例において、以下のタイムスタンプパターンが使用される。

タイムスタンプパターン 意味

\* 0以上の数字に一致するワイルドカードキャラクタ

20 . 1つの数字に一致するワイルドカードキャラクタ  
m-n mないしn

, または

(.) キャラクタグループ

— 2つの日付 (絶対的または相対の日付) の範囲を指定する日付範囲指定子

【0017】上記のタイムスタンプパターンは、たとえば、年、月、日、時間、分または秒がわからない場合にこれらを指定するために、URLのtime=またはrtime= (相対タイムスタンプ) フィールドにおいて使用される場合がある。たとえば、

## 3日におけるすべての

res.htmlページ

【0019】可変タイムスタンプを含むURL要求をサーバが受信すると、サーバは、クライアントが同一資源の異なるバージョンの一覧を要求していることを認識する。サーバは、アーカイブをすべて検索して、一致される資源をすべて識別し、すべての一致された資源をハイパーリンクにより示すHTMLページを返送する。ウェブ資源の中には多数のバージョンがアーカイブされているものもあるため、ユーザは、どのようにリンクを提示するかをオプションで指定することが可能である。したがって、本発明の別の特徴によれば、「タイムオーグ」パラメータにより、ユーザは、一致された資源に対応する

\* するリンクをどのように表示するかを指定することができる。たとえば、`timeorder=increase`は、昇順のタイムオーグでリンクを提示する。

【0020】加えて、これらのリンクは、案内および選択を容易にするためにカレンダーのようなフォーマットで提示することが可能である。たとえば、一致する資源のリンクが数年間に及ぶ場合、以下のように、すなわち、月番号に下線を引いて一致する資源の既存のバージョンを示しながら、リンクを表示することが可能である。

```

1996 1 2 3 4 5 6 7 8 9 10 11 12
1997 1 2 3 4 5 6 7 8 9 10 11 12
1998 1 2 3 4 5 6 7 8 9 10 11 12

```

1998年3月のリンクに対応するURLであれば、`http://www.a.com/res.html?time=1998_3_&timeorder=calendar`の形式を有する。また、タイムオーグは、時間単位について指定することも可能である。たとえば、`timeorder=+D`は、まず各月の1日目に対応するリンク、次に各月の2日目に対応するリンク等、以下同様に、一致する資源のリンクを昇順の日付順で列記することを意味する。同様に、`timeorder=+D`は、一致する資源のリンクを昇順の日付順、次に昇順の年号順に列記することを意味する。

【0021】相対タイムスタンプは、時間値の最後に“”ワイルドカードを用いることで拡張することが可能である。たとえば、現行日が1998年12月9日とすると、`rtime=next_month*`は、`rtime=1999_1_*`（1999年1月の任意の日）を意味する。同様に、現行日が1998年12月9日とすると、`rtime=+1y*`は、`rtime=1999_12_9_*`（1999年12月9日の任意の時間）を意味する。

【0022】本明細書において「要求ヘッダ方式」と称する代替の実施例において、タイムスタンプは、タイムスタンプ：1998年6月9日  
等HTTP要求ヘッダの一つとして表示されることが可能である。本明細書において「特殊キヤラクタ方式」と称する別の実施形態において、特殊キヤラクタは、`http://www.cnnfn.com;time=2+2+1998&timezone=server`等のURLにタイムスタンプが含まれていることを示すために使用されることが可能である。

【0023】（ウェブブラウザ）ウェブブラウザ100は、本発明の特徴および機能を適用するように本明細書において変形される。Microsoft Internet Explorer™、Netscape 50

avigator™等従来のブラウザとして具現化されてもよい。詳細を後述するように、ウェブブラウザ100は、ユーザが希望の日時を指定することができるように新たなオプション選択パネルを組み込むだけでよい。実際、ユーザがブラウザ100の「探索」ウィンドウにおいて表示されるURLにタイムスタンプを手動で添付して、従来のウェブブラウザ100を使用することが可能である。

【0024】一実施例において、ユーザは、タイムスタンプをONまたはOFFするオプションを有する。タイムスタンプが活性化されている場合、ブラウザ100は、これに応じてURLを変更した後、URLをウェブ130へ送信する。対応するウェブサーバ140、150がタイムスタンプを認識できる保証はないため、サーバ140、150により返送された文書には、タイムスタンプを含まない埋め込み式ハイパーリンクが含まれることがある。このため、この状況において、ウェブブラウザ100は、埋め込み式ハイパーリンクに関連付けられたURLを自動的に変換して、タイムスタンプオプションが活性化されている場合にユーザがハイパーリンクをクリックすると適当なタイムスタンプを追加することが可能である。ウェブブラウザ100は、選択されたタイムスタンプフォーマットにしたがってURLを変換すればよい。要求ヘッダ方式の実施例において、特殊な要求ヘッダ“Time-Stamp: June 9, 1998”（「タイムスタンプ：1998年6月9日」）を送信するようにブラウザ100を変形すればよい。さらに、特殊なタイムスタンプを有するいずれの埋め込み式ハイパーリンクについても新たなタイムスタンプタグを含むようにHTMLを変形すればよい。たとえば、`<A HREF="www.lucent.com">Lucent WebSite</A>`等のハイパーリンクについて、HTMLは、以下のよう

に1998年2月2日のタイムスタンプを示すように変形すればよい。

```
<A HREF="www.lucnet.com">
<TIMESTAMPTIME="2+2+1998"
TIMEZONE="server"></TIMESTAMP>
Lucnet Web Site</A>
```

【0025】(恒久ウェブサーバ)恒久ウェブサーバ140、150は、本明細書において後述する機能および動作を実行するように変形された従来のハードウェアおよびソフトウェアとして具体化される。すなわち、恒久ウェブサーバ140、150は、(i)タイムスタンプまたは相対タイムスタンプを含むURLを受信し、(ii)タイムスタンプを抽出し、(iii)適当なタイムスタンプに対応するウェブページを検索し、(iv)要求されたページをクライアントに返信する方法を把握する必要がある。恒久ウェブサーバ140、150は、抽出されたURLを、選択されたタイムスタンプフォーマットにしたがって解読すべきである。

【0026】要求された時間に対応するウェブ資源のバージョンが存在しない場合、本発明は、要求されたターゲット時間近傍で時間軸で記憶された文書のバージョンを提供する。たとえば、本発明では、ウェブ資源がアーカイブされた前のバージョンから変更されておらず、要求された時間から最近のタイムスタンプを有するウェブ資源のバージョンが提供されるとする。あるいは、要求された時間のすぐ次のタイムスタンプを有するウェブ資源のバージョンが提供される。

【0027】さらに、恒久ウェブサーバ140、150は、そのウェブを提供する履歴における情報をすべて保存する必要がある。したがって図1に示すように、サーバ140、150等それぞれの恒久ウェブサーバは、ウェブユーザにとって恒久的に利用可能なウェブ資源のバージョンをすべて記憶する恒久アーカイブ145、155をそれぞれ有する。恒久アーカイブ145、155は、いずれの記憶装置として具体化してもよいが、CD-ROM、CD-R、WORMまたはDVD-ROM等恒久的な(消去可能でない)記憶装置が好ましい。

【0028】恒久ウェブサーバ140、150が、日付の入ったURLに対応するためには、恒久ウェブサーバ140、150は、適時の情報の検索が可能となるようにそのコンテンツのすべてを時間順に記憶する必要がある。一実施例において、図2に示すように、恒久アーカイブ145、155は、ウェブサイトコンテンツ全体を、時系列のディレクトリ構造に関連してあるソートを行って、永久記憶装置に記憶する。図2は、ウェブサイトのコンテンツを時系列で配置したディレクトリ構造200を示す。したがって、リーフ210等、ディレクトリ構造200におけるそれぞれのリーフは、日付の入ったURLに対応する。たとえば、

<http://www.nytimes.com?ti> 50

me=24+2+1998

等日付の入ったURLは、  
<http://www.nytimes.com/archive/1998/2/24/>  
 に概念的に等しい。

【0029】記憶装置の用途において、ウェブサイトコンテンツ全体を記憶することは非効率的であることは勿論である。多くのウェブページにおいては、日々の変更はたとえあつたとしてもごくわずかである。したがって、アーカイブにおける冗長を単に除去することによって、十分な記憶効率を得られる。冗長が除去されると、通常のウェブサイト記憶容量に加えての記憶要件は、一般にさほど大きくない。

【0030】図3は、恒久アーカイブ145、155の冗長を低減するアーカイブプロセス300を示す。当該アルゴリズムにおいて記憶されるすべてのファイルまたはサブディレクトリは、図2の例示的なディレクトリ構造200のアーカイブサブディレクトリ220の下にある。図3に示すように、アーカイブプロセス300は、まず、ステップ310において、サブディレクトリA等それぞれのサブディレクトリについて、サブディレクトリA以前に作成され、かつこれと同一コンテンツを有するサブディレクトリBが存在するかを判定するテストを実行する。ステップ310において、サブディレクトリA以前に作成され、かつこれと同一コンテンツを有するサブディレクトリBが存在しないと判定されると、恒久アーカイブ145、155のサブディレクトリレベルで冗長を低減することは可能でなく、プログラム制御はステップ330に進む。

【0031】しかしながら、ステップ310において、サブディレクトリA以前に作成され、かつこれと同一コンテンツを有するサブディレクトリBが存在すると判定されると、サブディレクトリAは、ステップ320において、サブディレクトリBを示すエイリアスとなる。たとえば、図4(a)に示すように、ウェブサイトの現行コンテンツが前日のコンテンツと同一である場合、昨日のサブディレクトリを示す今日のエイリアスが作成される。同様に、図4(b)に示すように、現行日のコンテンツが前月のコンテンツと同一である場合、先月のサブディレクトリを示す今月のエイリアスが作成される。

【0032】その後、ステップ330において、ファイルA等それぞれのファイルについて、ファイルA以前に作成され、かつこれと同一コンテンツを有するファイルBが存在するかを判定するテストを実行する。ステップ330において、ファイルA以前に作成され、かつこれと同一コンテンツを有するファイルBが存在しないと判定されると、ファイルレベルで恒久アーカイブ145、155の冗長を低減することは可能でない。このため、プログラム制御はステップ350で終了する。

【0033】しかしながら、ステップ330において、



ファイルA以前に作成され、かつこれと同一コンテンツを有するファイルBが存在すると判定されると、ファイルAは、ステップ340において、ファイルBを示すエイリアスとなる。その後、プログラム制御は、ステップ350において終了する。

【0034】アーカイブプロセス300は、一致ファイルまたはディレクトリを検索する必要があるため実用的ではない場合がある。アーカイブにおけるエンティティの数とともにランタイムが急激に増大する。多数の次善策が可能であることは、当業者にとって明らかであろう。非常に単純な解決策としては、最も最近追加されたアーカイブ（昨日のコンテンツのように）に対して今日アーカイブしようとするものを単にチェックするものがある。ウェブサイトのほとんどは、最も最近追加されたものごとく僅かしか異なっていないため、この方法はかなり妥当である。この方法は、ファイルシステムの周知のインクリメンタルバックアップと同様である。

【0035】ウェブサーバが恒久的でない場合、影響はごく僅かである。一実施形態において、要求が、ウェブサーバにより記録されないタイムスタンプを含む場合、このサーバは、要求されたウェブ資源の最近のバージョンを配信するようにするといよい。

【0036】恒久アーカイブの記憶要件を低減する別の方法として、アーカイブされた正しいデータの検索についてウェブサーバをよりスマートにするものがある。たとえば、ウェブ資源の恒久的記憶は、ウェブ資源のすでに保存されているバージョンに対して多少異なるバージョンに限定されるようにしてもよい。たとえば、例示のアーカイブが、6/4/1996、6/12/1996、6/23/1997、2/1/1998、および2/3/1998の5つの異なるウェブ資源のバージョンを有する場合、ウェブサーバは、要求された日付がアーカイブされているバージョンのいずれとも等しくない場合に、要求された日付が最も近い過去の日付を有するバージョンと同一であるとする。さらに、ディレクトリにおいて特殊なシンボリック（または、Mac OS（登録商標）におけるエイリアス、MS Windows（登録商標）におけるショートカット）を用いることで、現行ディレクトリの下で見つからなかったファイルまたはディレクトリをどこで探すかを表すことができる。このように、変更された部分のみが適当なディレクトリの下で記憶される。変更されなかったデータはすべてかかる特殊なリンクのチェーンを介して照会することが可能である。

【0037】(DNSサーバ)ドメイン名サーバ(DNS)は、以下に記載する機能および動作を行うように本明細書において変形されるような、従来のハードウェアおよびソフトウェアとして具体化されてもよい。従来のDNSサーバは、DNSデータベースにないドメイン名照会を拒絶する。本発明に係る日付の入ったURLの利

点の1つは、これを用いて履歴としてのウェブ資源を照会することができることである。たとえば、A社がB社に吸収される場合、“www. A. com”を通して照会されるすべてのウェブページが有効でなくなる場合がある。A社からのいくつかの文書にアクセスしたいユーザにとって、参照をすべてB社のウェブサイトにどこかに変更する必要がある。

【0038】DNSサーバが、名称照会を拒絶しない代わりに、A社の履歴情報が所在する場所を知っているアーカイブサービス会社に照会する場合には、A社の履歴情報に依然としてアクセスすることができる。DNSサーバ自体は、より素早く名称をIPアドレスプロセスに分解するために履歴データをいくつか記憶することも可能である。図5(a)および(b)は、それぞれA社とB社の合併の前後にDNSサーバデータベースに記憶されたデータの例である。図5(b)に示すように、ユーザが、合併後www. A. comを見つけた場合、DNSサーバは、B社と関連づけられた新たなIPアドレスにユーザの要求をリダイレクトするに十分な情報を有する。データベースに列記された日付は、対応するドメイン名の有効期間である。したがって、“www. A. com 2/2/1999”のような日付の入ったドメイン名照会は無効であるが、“www. A. com 2/2/1992”は有効である。

【0039】図6は、本発明に係るDNSサーバプロセス600を示す。図6に示すように、DNSサーバプロセス600は、まず、ステップ610においてドメイン名要求を受信する。ステップ620において、このドメイン名要求が日付つきであるかを判定するテストを実行する。ステップ620において、ドメイン名要求が日付つきでない判定されると、ステップ630において通常の名称検索結果を返信する。

【0040】しかしながら、ステップ620において、ドメイン名要求が日付つきであると判定されると、DNSサーバプロセス600は、ステップ640において、DNSデータベースから日付制約のあるドメイン名を検索する。ステップ650において、日付の入ったドメイン名が見つかったかを判定するさらなるテストを実行する。ステップ650において、日付の入ったドメイン名が見つからなかったと判定されると、ステップ660において、DNSサーバはアーカイブサービス会社に照会してさらに検索した後、プログラム制御はステップ670に進む。

【0041】しかしながら、ステップ650において、日付の入ったドメイン名が見つかったと判定されると、ステップ670において、リダイレクトの場合に検索結果および表示が返信され、プログラム制御は終了する。

【0042】ドメイン名がDNSサーバにより分解されると、ウェブブラウザ100は、DNSサーバから返信されたものにしたがって要求をウェブサーバ140、1

50に送信する必要がある。たとえば、ユーザからの  
 “http://www.A.com?time=2+2+1992”というURLの要求により、ブラウザ100は、“www.A.com 2/2/1992”等のフォーマットでドメイン名分解要求をDNSサーバに送信する。ここで、A社はB社の一部であるため、その結果は“123.2.3.222 redirect”のようになる。これでウェブブラウザ100は、サーバのIPアドレスを有し、これがリダイレクトであることもわかる。このため、ウェブブラウザ100は、実際には、“http://123.2.3.222?http://www.A.com&time=2+2+1992”

URL	開始日	終了日	IP アドレス	リダイレクト URL
WWW.A.COM	13/2/1980	1/5/1998	123.2.3.222	WWW.B.COM
WWW.A.COM	1/1/2000	現在	234.2.2.12	—
WWW.B.COM	23/1/1985	現在	123.2.3.222	—

【0044】このように、ドメイン名は、これを永久に無効にすることなく再使用することが可能である。

【0045】（動的ウェブコンテンツに対する恒久的アクセス）現在ウェブは、たとえばWeb Camからのリアルタイム映像、たとえばWeb Cast イベントからの音声ストリーム、ならびにJava、JavaScriptまたはActive-X対応ウェブページ等の動的コンテンツであふれている。アプリケーションによっては、連続した動的コンテンツをアーカイブすることが必要でないまたは望ましくない場合もある。さらに、ウェブページの宣伝部分すべてをアーカイブすることが必要でないまたは望ましくない場合もある。

【0046】金融または株式ブローカー取引等の時系列データの場、所定のタイムスタンプ制限について記録の部分を抽出することは容易である。たとえば、所定の日座残高を確認するために、ブラウザ100は、  
 http://bank.com/?id=12345&time=2+3+1998  
 の形式で要求を送信することができる。

【0047】サーバ140、150は、1998年3月2日以前のデータを検索または再計算して結果を返信するだけでよい。いずれにしてもかかるアプリケーション環境におけるすべての取引がタイムスタンプを有しているため、この機能をサービスに追加することは簡単である。

【0048】リアルタイムコンテンツの場合、タイムスタンプを添付する際の唯一の制約は記憶要件である。アーカイブすべき情報量に比べて使用可能な記憶スペースが多い場合、ウェブサイト管理者は、リアルタイムコンテンツをアーカイブするか、あるいは1日、1週間または1年分のデータ等、その一部をアーカイブするように選択することができる。

\*92”等の形式で、要求をB社のウェブサーバ140、150に送信する。B社のウェブサーバ140、150は、このA社の旧アドレスを適当な場所にどのようにマッピングし、正しい情報を入力するかわかる。

【0043】この新たなDNSサーバの1つの副次的利点は、いくつかの名称は、履歴になると再使用できることである。たとえば、A社という名称の別の会社は、以下のエントリを有するDNSデータベースを更新することによって、所定の期間経過後にwww.A.comドメイン名を使用することが可能である。

【表1】

【0049】動的に作成された宣伝において、ウェブサイト管理者は、昔の宣伝を（何らかの特別の理由により）「再度表示」することが妥当であるか、または昔の宣伝を、「リアルタイムで」アーカイブされたウェブコンテンツに関係のない新しい最新のコマースに替えることができるかを決定しなければならない。

【0050】ウェブページ上で動的に実行されたスクリプトの場合、2つの状況の可能性があり、スクリプトがいずれの時間間隔にも関連づけられない場合、スクリプトは安全にアーカイブされる。しかしながら、スクリプトが、クロック、カレンダー表示等何らかの時間間隔に付随してその挙動を変更する場合、スクリプトは、日付の入ったURLを処理できるように修正する必要がある。Java、JavaScriptまたはActive-X規格に対する修正が適宜必要となる。CGIアプリケーションにより動的に作成されるウェブページも同様に処理される。

【0051】（インターネットアーカイブサービス）本発明は、すべてのウェブサーバがURLにおけるタイムスタンプを直接処理する必要なく、恒久的なウェブアクセスを提供する。一般に、ユーザは、たとえば、ウェブブラウザにより日付の入ったURLを上記のように入力する。ブラウザは、日付の入ったURLを適切なウェブサーバに送信する。上述したように、開示されるウェブブラウザは、所望の日時の指定を容易にする機構をオプションで有することが可能であり、あるいは、ユーザは、ブラウザの「探索」ウィンドウにおいて表示されるURLに手動でタイムスタンプを添付することができる。ウェブサーバは、URL（タイムスタンプ、相対タイムスタンプまたは可変タイムスタンプを含む）を受信する。そして、ウェブサーバは、その自アーカイブから正しいウェブページ（または複数のウェブページ）を検

17

業し、または本発明にしたがって所定のインターネットアーカイブサイトにユーザをリダイレクトする。

【0052】たとえば、ユーザは、以下の日付の入ったURLを入力して1999年の大晦日のニュースウェブサイトのメインページにアクセスする。

http://www.news-site.com/?rtime=1999\_12\_31

【0053】本発明によれば、ニュースウェブサイトは、ユーザ要求を以下のURLにリダイレクトする。

http://www.archive.com/news-site.com?rtime=1999\_12\_31

【0054】ニュースウェブサイトのアーカイブをホストするインターネットアーカイブサービスは、リダイレクトされたURLがニュースウェブサイトからのコンテンツのものであることがわかるので、適当な資源を検索する。

【0055】図7(a)は、要求された資源がアーカイブされている対応インターネットアーカイブサービス160をウェブサーバ140が把握している実施形態において、ユーザのブラウザ100、ウェブサーバ140およびインターネットアーカイブサービス160の間の各種通信を示す。図7(a)に示すように、ユーザは、まず、(ブラウザ100により)要求710をウェブサーバ140に送信する。ウェブサーバ140は、この要求がアーカイブされている資源に対するものであることを認識し、URLリダイレクト応答720を返信する。ブラウザ100は、上述した以下の例等のURLリダイレクト応答720を用いて、メッセージ730によりインターネットアーカイブサービス160に連絡する。  
http://www.archive.com/news-site.com?rtime=1999\_12\_31

インターネットアーカイブサービス160は、日付の入ったURLを用いて、所望の資源を検索する、またはインターネットクエリの場合には対応する資源のリンクのページを形成し、メッセージ740においてブラウザ100に返信する。

【0056】図7(b)は、要求された資源がアーカイブされている対応インターネットアーカイブサービス160をウェブサーバ140が把握していない実施形態において、ユーザのブラウザ100、ウェブサーバ140、アーカイブファインダ750およびインターネットアーカイブサービス160の間の各種通信を示す。図7(b)に示すように、ユーザは、まず、(ブラウザ100により)要求755をウェブサーバ140に送信する。ウェブサーバ140は、この要求がアーカイブされている資源に対するものであることを認識し、要求760をアーカイブファインダ750に送信して、要求された資源がアーカイブされている特定のインターネットア

18

ーカイブサービス160を識別する。アーカイブファインダ750は、要求を評価し、サイトロード、サイトの位置、要求の性質等、各種基準にしたがって要求について最良のインターネットアーカイブサービス160を見つけ出す。アーカイブファインダ750は、メッセージ765において、最良のインターネットアーカイブサービス160が一致したことをウェブサーバ140に返信する。ウェブサーバ140は、次に、識別された最良のインターネットアーカイブサービス160をブラウザ100にURLリダイレクト応答770として送り出す。ブラウザ100は、上述した以下の例等のURLリダイレクト応答770を用いて、メッセージ775により、識別されたインターネットアーカイブサービス160に連絡する。

http://www.archive.com/news-site.com?rtime=1999\_12\_31

識別されたインターネットアーカイブサービス160は、日付の入ったURLを用いて、所望の資源を検索する、またはインターネットクエリの場合には対応する資源のリンクのページを形成し、メッセージ780においてブラウザ100に返信する。

【0057】本明細書に図示および記載の実施形態および変形例は本発明の原理を示す単なる例示であり、本発明の範囲および精神から逸脱することなく当業者により各種変形が実施可能であることが理解されるはずである。

【図面の簡単な説明】

【図1】本発明に係るインターネットまたはワールドワイドウェブ(「ウェブ」)環境を示す図である。

【図2】ウェブサイトのコンテンツを時系列で配置したディレクトリ構造を示す図である。

【図3】図1の恒久アーカイブの冗長を低減するためにウェブサイトに実施されるアーカイブプロセスの例を示すフローチャートである。

【図4a】図3のアーカイブプロセスによりすでにアーカイブされているコンテンツと同一のウェブサイトのコンテンツのエリアシングを示す図である。

【図4b】図3のアーカイブプロセスによりすでにアーカイブされているコンテンツと同一のウェブサイトのコンテンツのエリアシングを示す図である。

【図5a】本発明に係るDNSサーバデータベースのサンプル表である。

【図5b】本発明に係るDNSサーバデータベースのサンプル表である。

【図6】タイムスタンプのついたURLに対応するIPアドレスを決定するためにDNSサーバにより実行されるDNSサーバプロセスの例を説明するフローチャートである。

【図7a】本発明の2つの実施形態に係る、図1のユー

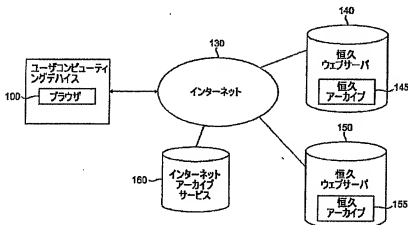
ザのブラウザ100、ウェブサーバ140およびインターネットアーカイブサービス160の間の各種通信を示す図である。

【図7b】本発明の2つの実施形態に係る、図1のユーザのブラウザ100、ウェブサーバ140およびインターネットアーカイブサービス160の間の各種通信を示す図である。

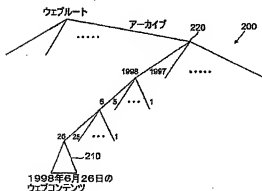
\*【符号の説明】

- 100 ブラウザ
- 130 インターネット
- 140、150 恒久ウェブサーバ
- 145、155 恒久アーカイブ
- 160 インターネットアーカイブサービス

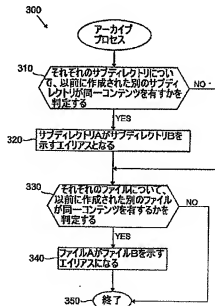
【図1】



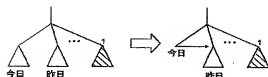
【図2】



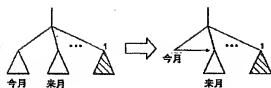
【図3】



【図4a】



【図 4 b】



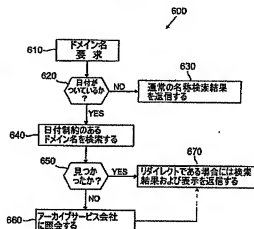
【図 5 a】

URL	開始日	終了日	IP アドレス	リダイレクト
WWW.A.COM	1997/10/01	現在	222.1.23.22	-
WWW.B.COM	2001/1/01	現在	222.2.3.222	-

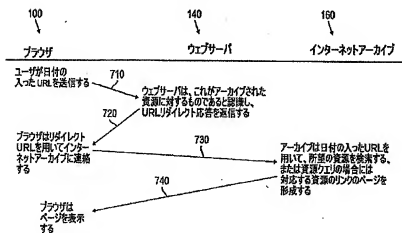
【図 5 b】

URL	開始日	終了日	IP アドレス	リダイレクト
WWW.A.COM	13/01/98	18/1/98	123.2.3.222	WWW.B.COM
WWW.B.COM	29/1/98	現在	123.2.3.222	-

【図 6】



【図 7 a】



【図7b】

